

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

J1040 U.S. PRO  
10/077966  
02/20/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 9月27日

出 願 番 号

Application Number:

特願2001-295397

[ST.10/C]:

[JP2001-295397]

出 願 人

Applicant(s):

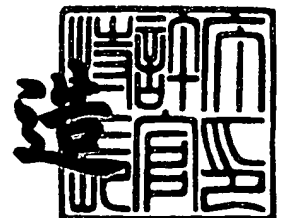
株式会社日立製作所

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2002年 2月 1日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2002-3003176

520.41229X00



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): NAKAMURA, et al.  
Serial No.: Not yet assigned  
Filed: February 20, 2002  
Title: DATA STORAGE SYSTEM AND CONTROL METHOD THEREOF  
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of  
Patents and Trademarks  
Washington, D.C. 20231

February 20, 2002

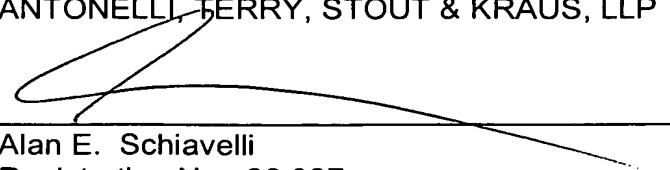
Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2001-295397, filed September 27, 2001.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Alan E. Schiavelli  
Registration No. 32,087

AES/alb  
Attachment  
(703) 312-6600

【書類名】 特許願

【整理番号】 K01011061A

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/00

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 中村 泰明

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 中野 俊夫

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 島田 朗伸

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 村上 達也

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

【氏名】 守島 浩

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9902691

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データストレージシステム及びこの制御方法

【特許請求の範囲】

【請求項 1】

情報を格納する複数の外部記憶装置が第 1 のネットワークに接続され、それぞれが分離して配置されたデータストレージシステムの管理方法において、

管理サーバに対して外部記憶装置から割り込みがかかる第 1 のステップと、

前記管理サーバから外部記憶装置に対して排他制御用コマンドを発行する第 2 のステップと、

前記外部記憶装置が応答した構成情報を、前記管理サーバが受領する第 3 のステップと、

前記管理サーバが受領した構成情報を、データベースに格納する第 4 のステップ

とを有するデータストレージシステムの制御方法。

【請求項 2】

情報を格納する複数の外部記憶装置が第 1 のネットワークに接続され、それぞれが分離して配置されたデータストレージシステムの管理方法において、

管理サーバから複数の外部記憶装置に対して排他制御用コマンドを発行する第 1 のステップと、

前記外部記憶装置が応答した構成情報を、前記管理サーバが受領する第 2 のステップと、

前記管理サーバが受領した構成情報を、データベースに格納する第 3 のステップ

とを有するデータストレージシステムの制御方法。

【請求項 3】

請求項 2 記載のデータストレージシステムの制御方法において、

前記排他制御コマンドにより、前記管理サーバが、前記外部記憶装置全体の構成情報を、系時的に取得し、当該管理サーバの前記データベースに格納するデータストレージシステムの制御方法。

【請求項4】

請求項3記載のデータストレージシステムの制御方法において、  
前記系時的な取得は、複数の外部記憶装置に対する同時かつ定期的な問合せを契機とするデータストレージシステムの制御方法。

【請求項5】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置とが、それぞれ、ネットワークに接続され、それぞれが分離して配置されたデータストレージシステムの管理方法において、

管理サーバに対してログインして、アクセス許可を要求する第1のステップと

前記管理サーバから外部記憶装置に対して排他制御用コマンドを発行する第2のステップと、

前記外部記憶装置が応答した構成情報を、前記管理サーバが受領する第3のステップと、

前記管理サーバが受領した構成情報を、データベースに格納する第4のステップ

とを有するデータストレージシステムの制御方法。

【請求項6】

請求項5記載のデータストレージシステムの制御方法において、

前記排他制御コマンドにより、前記管理サーバが、前記外部記憶装置全体の構成情報を取得し、当該管理サーバの前記データベースに格納するデータストレージシステムの制御方法。

【請求項7】

請求項5記載のデータストレージシステムの制御方法において、更に、

第2のステップの排他制御用コマンドの発行に基づいて、前記複数のコンピュータのアプリケーションプログラムを起動する第5のステップと、

前記管理サーバが前記複数のコンピュータからホスト論理構成情報を受領する第6のステップ

とを有するデータストレージシステムの制御方法。

【請求項 8】

請求項 7 記載のデータストレージシステムの制御方法において、

第 4 のステップで格納される構成情報と、前記ホスト論理構成情報とが対応付けられてデータベースに格納されるデータストレージシステムの制御方法。

【請求項 9】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置とが、それぞれ、ネットワークに接続され、それぞれが分離して配置されたデータストレージシステムの管理方法において、

管理サーバに対してログインして、アクセス許可を要求する第 1 のステップと

前記管理サーバから構成情報を送出する第 2 のステップと、

前記管理サーバに対し、構成情報の変更を指示する第 3 のステップと、

前記管理サーバから複数の外部記憶装置に対して排他制御用コマンドを発行する第 4 のステップと、

前記外部記憶装置が応答した構成情報の設定の完了を、前記管理サーバが受領する第 5 のステップと、

前記管理サーバが構成情報の変更を、データベースに格納する第 6 のステップとを有するデータストレージシステムの制御方法。

【請求項 10】

請求項 6 又は 9 記載のデータストレージシステムの制御方法において、前記管理サーバが取り扱う構成情報には、

外部記憶装置の内部アクセスパス、論理ディスク、論理ディスクの容量、論理ディスクへのアクセス権限、データ移動に関する設定、

外部記憶装置間のデータコピーに関する設定、

外部記憶装置の性能制御モード若しくは性能データの設定又は取得、

データストレージシステム保守方法の設定又は障害発生若しくは障害通知、又は、

ユーザによる操作

が含まれることを特徴とするデータストレージシステムの制御方法。

【請求項 1 1】

請求項 9 記載のデータストレージシステムの制御方法において、

前記構成情報の変更の対象である外部記憶装置を識別して、第 4 のステップの排他制御用コマンドを当該外部記憶装置にのみ発行するデータストレージシステムの制御方法。

【請求項 1 2】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置とが、それぞれ、ネットワークに接続され、それぞれが分離して配置されたデータストレージシステムの管理方法において、

管理サーバから複数の外部記憶装置に対して排他制御用コマンドを発行する第 1 のステップと、

前記外部記憶装置が応答した構成情報を、前記管理サーバが受領する第 2 のステップと、

前記排他制御用コマンドの発行に基づいて、前記複数のコンピュータのアプリケーションプログラムを起動する第 3 のステップと、

前記管理サーバが前記複数のコンピュータからホスト論理構成情報を受領する第 4 のステップと

前記管理サーバが受領した構成情報と、前記ホスト論理構成情報とを、データベースに格納する第 5 のステップ

とを有するデータストレージシステムの制御方法。

【請求項 1 3】

請求項 1 2 記載のデータストレージシステムの制御方法において、更に、

前記管理サーバに対して外部記憶装置から割り込みがかかる第 6 のステップを有するデータストレージシステムの制御方法。

【請求項 1 4】

請求項 1 2 又は 1 3 記載のデータストレージシステムの制御方法において、

前記排他制御コマンドにより、前記管理サーバが、前記データストレージシステム全体の構成情報を、系時的に取得し、対応付けて、当該管理サーバの前記データベースに格納するデータストレージシステムの制御方法。



【請求項 1 5】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置とが、それぞれ、ネットワークに接続され、それぞれが分離して配置され、第 1 のネットワークを介して接続された管理サーバを有するデータストレージシステムの管理方法において、

管理サーバに、前記コンピュータが使用するファイル種別、時刻を入力する第 1 のステップと、

前記管理サーバが、構成情報データベースに基づいて、前記ファイルに対応する論理ディスクの物理的な格納位置を検索して表示する第 2 のステップと、

前記物理的な格納位置に関連する別の論理ディスクを検索して表示する第 3 のステップと、

データストレージシステムの変更の履歴を蓄積しているデータの中から、前記時刻以前で前記格納位置に関連したシステムの変更の内容を検索し表示する第 4 のステップと、

論理ディスクの性能の履歴を蓄積しているデータの中から、前記時刻以後の論理ボリュームの性能を検索し表示する第 5 のステップと、

前記論理ボリュームの性能が低下しているときに、前記システムの変更の内容を表示又は通知する第 6 のステップ

とを有するデータストレージシステムの制御方法。

【請求項 1 6】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置とが、それぞれ、ネットワークに接続され、それぞれが分離して配置され、第 1 のネットワークを介して接続された管理サーバを有するデータストレージシステムの管理方法において、

管理サーバが、前記コンピュータのアプリケーションソフトが使用するファイルの容量を、当該コンピュータに問い合わせ、系時的に応答を受領する第 1 のステップと、

管理サーバが、前記ファイルを格納した論理ディスクとの対応付けを、構成情報データベースの内容から検索し、当該論理ディスクの容量と、前記系時的なフ

ファイルの容量との関係を示す第 2 のステップ  
とを有するデータストレージシステムの制御方法。

【請求項 1 7】

請求項 1 6 記載のデータストレージシステムの制御方法において、  
前記関係を示す第 2 のステップは、構成情報データベースの内容から、前記論理ディスクの容量と前記ファイル容量が等しくなる時刻を予測して、表示又は通知することであるデータストレージシステムの制御方法。

【請求項 1 8】

情報を格納する複数の外部記憶装置が、ネットワークに接続され、それぞれが分離して配置されたデータストレージシステムであって、

各外部記憶装置は、自己の構成を定義し、参照し、性能やデータを提示し、又は、障害を知らせるため、イベント情報を送出する外部接続用インタフェースを持ち、

前記データストレージシステムは、その一部に管理サーバ部を有し、

当該管理サーバ部は、前記外部記憶装置と接続され、前記外部接続用インタフェースを介して、当該複数の外部記憶装置のイベント情報を、系時的に蓄積する構成情報データベースを有すること

を特徴とするデータストレージシステム。

【請求項 1 9】

請求項 1 8 記載のデータストレージシステムにおいて、

前記構成情報データベースに系時的に蓄積する際には、前記管理サーバ部は、排他制御用コマンドを前記複数の外部記憶装置に対して発行するデータストレージシステム。

【請求項 2 0】

情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置が、ネットワークに接続され、それぞれが分離して配置されたデータストレージシステムであって、

各コンピュータは、自己のホスト論理構成情報を取得するためのアプリケーションを搭載しており、

各外部記憶装置は、自己の構成を定義し、参照し、性能やデータを提示し、障害を知らせるため、イベント情報を送出する外部接続用インタフェースを持ち、前記データストレージシステムは、その一部に管理サーバ部を有し、当該管理サーバ部は、

前記外部記憶装置と接続され、前記外部接続用インタフェースを介して、当該複数の外部記憶装置のイベント情報、及び

前記コンピュータと接続され、前記ネットワークを介して、当該複数のコンピュータのホスト論理構成情報を、

系時的に蓄積する構成情報データベースを有すること  
を特徴とするデータストレージシステム。

【請求項 21】

請求項 20 記載のデータストレージシステムにおいて、

前記構成情報データベースに系時的に蓄積するに際し、前記管理サーバ部は、前記複数の外部記憶装置のイベント情報と、前記複数のコンピュータのホスト論理構成情報とを、対応させるデータストレージシステム。

【請求項 22】

請求項 19 又は 21 記載のデータストレージシステムにおいて、

前記管理サーバは、前記コンピュータが扱うファイルと時刻情報を特定することで、前記構成情報データベースを検索する機能を有するデータストレージシステム。

【請求項 23】

請求項 19 又は 21 記載のデータストレージシステムにおいて、

前記管理サーバは、システム構成の変更の履歴、又は、システムの性能の履歴を表示する機能を有するデータストレージシステム。

【請求項 24】

請求項 19 又は 21 記載のデータストレージシステムにおいて、

前記管理サーバは、前記コンピュータのアプリケーションが使用するファイルの容量が、前記外部記憶装置の論理ディスクの容量に到達する時期を通知する機能を有するデータストレージシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、情報を利用する複数のコンピュータと、情報を格納する複数の外部記憶装置（以下、ディスクサブシステムという。）とが、ネットワークに接続され、分離して配置された場合のデータストレージシステム（以下、単にストレージシステムという。）の集中管理技術に係り、特に、複数のディスクサブシステム間にまたがったストレージシステム全体の管理技術に関する。

【0002】

【従来の技術】

複数の外部記憶装置、例えば、複数のディスクサブシステムにまたがったデータの集中管理を行うには、各システムの構成情報を取得し、システム全体を統括した全体構成を定義する必要がある。ここで、構成情報には、例えば、ディスクサブシステムの内部アクセスパス、論理ディスク、論理ディスクの容量やアクセス権限、データ移動に関する設定、ディスクサブシステム間のデータコピーに関する設定、性能制御モードや性能データの設定、取得、保守方法の設定及び障害やユーザ操作によるイベントなどがある。

【0003】

従来は、システム管理者らが、ディスクサブシステムを利用するホストコンピュータ（以下、単にホストと称する）が管理するソフトウェアを用いて、ディスクサブシステムの構成や性能、障害、増設その他システム配下で発生する事象（以下、イベントと称する）を、定期的に収集していた。つまり、システム管理者が、マニュアル操作で管理ソフトウェアを用いて、個々のディスクサブシステムに接続し、これらのシステムの構成情報を取得し、システムの全体構成の定義や必要な設定をせざるをえなかった。

【0004】

米国特許第5,973,690号では、ホストから1つのディスクサブシステムに対してアクセス可能な論理ボリュームが、ディスクサブシステムのどの物理ディスクに対応しているかをマッピング表示するための技術が開示されている。し

かし、複数のディスクサブシステムの間を横断的に管理することは示唆が無い。

【0005】

システム全体を統括した全体構成を定義するには、システム管理者が、複数のディスクサブシステム間にまたがった設定を、一括して行えることが望ましい。システム全体の構成を定義する際に、ディスクサブシステム毎に行うよりも容易で、かつ、構成を調べ定義し直す回数が減ることで、人為的なミスオペレーションが低減されるからである。複数のディスクサブシステム間にまたがった設定を一括して行えば、システム運用の向上も図ることが可能だからである。

【0006】

あるユーザAがホストコンピュータにデータベースその他のアプリケーションAをインストールし、複数のディスクサブシステムを外部記憶装置として使用している状態を考える。ユーザAのアプリケーションAが使用するファイルの容量が手狭になったため、この外部記憶装置のシステム管理者Sが1つのディスクサブシステムに論理ディスク (LU: Logical Unit) を追加したとする。

【0007】

ところが、そのディスクサブシステムは、別のホストでユーザBが使用する別のアプリケーションB (アプリケーションAよりも、高性能を要求されるものとする) も使用している場合がある。

【0008】

このようなとき、追加された論理ディスクが、高性能を要求されるアプリケーションBが使用するために割り当てられた論理ディスクと、物理資源 (物理ディスク) を共有することとなっていた場合には、この論理ディスクの追加が影響し、性能重視のアプリケーションBの実行について、性能低下を引き起こすことになる。

【0009】

換言すると、システム管理者Sが行った、ユーザAのためにした論理ディスクの追加は、ユーザAのアプリケーションAの実行を維持増進するための措置ではあるが、アプリケーションBの実行性能の低下を引き起こし、システム全体の性

能面から見ると、結果として人為的なミスオペレーションであるといえる。

【 0 0 1 0 】

システム管理者が、アプリケーションの性能の監視を行うときには、通常、性能監視用のツールを用いて行う。性能監視用ツールは、アプリケーションのプロセス稼動状況や、アプリケーションが使用するファイルの読出と書込の性能を監視するものであるため、さきの論理ディスクの追加が、別のアプリケーションの性能低下を引き起こした原因であることを、突き止めることはできない。

【 0 0 1 1 】

インターネットの急速な普及に伴い、多数のクライアント端末からのアクセス要求が増加し、これらは、複数のホストからのアクセスとみなせる。これら多数のアクセスを統括するストレージシステムも、データ容量の需要に追従する措置が必要であり、個々のディスクサブシステムにおける論理ディスク増設の機会が恒常的に増加している。ビジネス関連のアプリケーションが使用するファイルに対応する論理ディスクが、いつ使用可能容量を超えるかを予測し、計画的な論理ディスク増設のスケジュールを立案することが望まれる。このためには、ファイル容量の増加傾向と、ファイルを格納した論理ディスクの位置と、その使用可能容量を調査し、これらの関係からスケジュールを立案しなくてはならない。

【 0 0 1 2 】

従来技術においては、これらの情報を個別に定期的に収集してはいたものの、これらの関係を構築する手段が無く、システム管理者の経験則に依存した措置に留まり、上記の予測、計画立案は非常に困難であった。

【 0 0 1 3 】

【発明が解決しようとする課題】

本発明は、複数（N台）のホストから共用される、複数（M台）のディスクサブシステムの構成において、複数のシステム管理者（L人）が、M台のディスクサブシステムを横断的に一括して管理し、所定の設定を迅速かつ簡易に実現するための管理技術を提供することを目的とする。

【 0 0 1 4 】

また、本発明は、ディスクサブシステムの構成の変更が、ホストで実行される

アプリケーションの性能に与える影響を把握可能な、ディスクサブシステムの管理技術を提供することを別の目的とする。

【 0 0 1 5 】

また、本発明は、ディスクサブシステムにおける計画的な論理ディスク容量の追加時期の決定が可能な管理技術を提供することを別の目的とする。

【 0 0 1 6 】

【課題を解決するための手段】

- 1) 一時的に全ての複数のディスクサブシステムへのアクセスを制限する排他制御コマンドを設ける。
- 2) この排他制御コマンドにより、複数のディスクサブシステム全体の構成情報、例えば、性能、設定の変更を、系時的に取得し、管理サーバ部のデータベース（“構成情報データベース”）に格納して、一元管理する。
- 3) アプリケーションが使用しているファイルの、論理ディスク上の位置を検出する機能を用いて、“構成情報データベース”に対応付けする機能を設ける。
- 4) “構成情報データベース”に、システム構成の変更内容や時刻をキーに検索できる手段を設ける。

【 0 0 1 7 】

【発明の実施の形態】

図 1 に、複数のホストコンピュータ 1 0 が、SAN（ストレージエリアネットワーク）4 0 を介して、データを送受し、共用する複数のディスクサブシステム 2 0 と接続される場合のシステム全体の概略構成を示す。各ディスクサブシステム 2 0 は、自己の構成を定義し、参照し、性能やデータを提示し、障害を知らせるため、イベント情報を送出するための外部接続用インタフェース 2 1 を持っている。

【 0 0 1 8 】

管理サーバ部 3 0 は、SAN 4 0 とは異なるローカルエリアネットワーク（LAN）5 0 のインタフェースで、複数のホスト 1 0 及び複数のディスクサブシステム 2 0 と接続できる。図 1 では管理サーバ 3 0 は 1 つのみ示されているが、複数あっても良い。また、管理サーバ 3 0 は、ディスクサブシステム 2 0 の内部に

設置されていても良いし、これら20から物理的に離れた場所に位置していても良い。「管理サーバ」なる用語には、サーバ機能を有する外部記憶装置の一部が含まれ、適宜、「管理サーバ部」と記す。

【0019】

ある1つの管理サーバ部30から、複数のディスクサブシステム20全体の構成を、これらの複数のサブシステム間にまたがって一括して定義することを考える。以下、単に、管理サーバ30と記す。

【0020】

管理サーバ30からシステム20へ、その管理サーバ30が唯一の設定手段となるように、排他制御用コマンドを発行する。ここで排他制御用コマンドとは、ある時間帯において任意に選択した複数のディスクサブシステム20を占有するコマンドをいう。占有時間は、長ければ1時間程度のこともあるが、別途、設定情報を作成しておき、わずかな占有時間で設定を行なう制御方法も用意されている。管理サーバ30は、また、その設定が正常終了するのを確認する機能を有する。

【0021】

管理サーバ30の機能ブロックを、図2を用いて説明する。

ユーザ管理レイヤ31は、管理サーバ30に接続する複数のユーザA～Cを管理する。ここでユーザには、システム管理者が含まれる。

オブジェクト管理レイヤ32は、各ディスクサブシステム20の構成情報の取得と、ユーザからの設定要求を管理する。

エージェント管理レイヤ33は、オブジェクト管理レイヤ32からの要求に応じ、排他制御コマンドを、サブシステムインタフェース341を介して、各ディスクサブシステム20へ発行する。

【0022】

インタフェースレイヤ34は、各ディスクサブシステム20とのデータの送受信を行なうサブシステムインタフェース341と、各ホストエージェント11とのアクセスを制御するホストインタフェース342を有している。

【0023】



排他制御が行なわれている間に、各ディスクサブシステム 20 の構成、性能、及び、障害その他のイベント情報を、オブジェクト管理レイヤ 32 が取得し、構成情報データベースに格納する。

【0024】

ユーザ管理レイヤ 31 によってアクセスを許可された唯一のシステム管理者（ユーザ）は、レイヤ 32 の構成情報データベースに格納された、複数のディスクサブシステム 20 の諸パラメータの変更、拡張、又は、削除を、同システム 20 にまたがり一括で行う。この結果、構成情報データベースと実際のディスクサブシステムの構成情報とが、所定の時点で異なることなく整合した状態とすることができる。

【0025】

管理サーバ 30 は、オブジェクト管理レイヤ 32 によって、システム 20 の構成の変更、拡張、削除を、自己の構成情報データベースへの登録も含め完了すると、占有していた複数のディスクサブシステム 20 全てを、エージェント管理レイヤ 33 によって開放する。

ここで前記のオブジェクト管理レイヤ 32 の構成情報データベースで、管理サーバ 30 が取り扱う情報は、各ディスクサブシステム 20 の内部アクセスパス、論理ディスク、これらの容量やアクセス権限、データ移動に関する設定、サブシステム間のデータコピーに関する設定、各ディスクサブシステムの性能・制御の設定、各ディスクサブシステムの性能データの取得、保守方法の設定及び障害やユーザオペレーションによるイベントの構成情報に関するものである。

【0026】

<諸情報の取得契機>

ディスクサブシステム 20 やホスト 10 の諸情報の取得契機は、唯一のシステム管理者（ユーザ）が管理サーバ 30 にアクセスして、管理サーバ 30 がシステム 20 の構成を定義する場合には、その構成をシステム 20 へ指示する前となる。この一方で、ディスクサブシステム 20 の障害、保守その他のイベントが発生した場合も取得契機と成り得る。具体的には、

- 1) 管理サーバ 30 により、各ディスクサブシステム 20 への定期的な問い合

わせにより、そのイベントを認識して情報を取得する場合である。

2) また、ディスクサブシステム20が検出した障害や保守のイベントが、サブシステムインタフェース341 (図2) からエージェント管理レイヤ33に通知された場合も、情報取得の契機となる。

【0027】

2) の場合、障害等のイベントを通知されたエージェント管理レイヤ33は、割り込み機能により、上位レイヤのオブジェクト管理レイヤ32に通知し、これを受けたオブジェクト管理レイヤ32により、管理サーバ30は、ディスクサブシステム20の状態が変更されたと認識する。このイベントを認識した後、管理サーバ30は、システム20の構成情報を取得し、構成情報データベースの情報を更新する。

【0028】

尚、自発的増設、障害や保守等のイベントによって、ディスクサブシステム20の構成が変更された場合に、管理サーバ30はその変更を特定し自己の構成情報データベースに登録することとなる。そこで管理サーバ30が所有する構成情報データベースの該当ディスクサブシステム20に、その旨のフラグを立てて管理すれば、そのデータベースの中でフラグの立っているシステム20のみを問い合わせ対象にすることで、それ以降の処理、特に、情報の取得を効率的に行なえる。

【0029】

#### <諸情報の取得方法>

上述の方法を用いて、複数のディスクサブシステムにまたがって、一括で所望の設定を行なう手順の一例を、図3を用いて説明する。複数のシステム管理者のうちのひとりが、複数のディスクサブシステム20 (ここでは2台) に対し、これらの構成を定義する際の手順である。この例では、ホストから論理ディスクへアクセス権限を与え、論理ディスクへの不当なアクセスを防止してデータの保護を行う、ディスクサブシステム固有の機能を用いている。

【0030】

1つの特定のホスト10 (図1) に接続された、2つのディスクサブシステム

XおよびYは、所定の数の論理ディスクを所有する。複数のホスト10が、複数のディスクサブシステム20を共用する環境下では、特定のホスト10がアクセスする論理ディスクに対し、他のホスト10からアクセスできないよう、セキュリティを設定する必要がある。

#### 【0031】

図3において、システム管理者Sは、管理サーバ30にログインし、アクセス許可を要求する（ステップ311）。これを受けて管理サーバ30は、ディスクサブシステムX及びYへ、その管理サーバ30がシステム全体の構成設定が可能な唯一の制御サーバとなるように、排他制御用コマンドを発行する（ステップ312）。

#### 【0032】

排他制御用コマンド発行の時点の各ディスクサブシステムX及びYの構成情報を、管理サーバ30が取得し（ステップ313）、構成情報データベースに格納する。

#### 【0033】

管理サーバ30にアクセス許可された唯一のシステム管理者Sは、構成情報データベースに格納されたディスクサブシステムX及びYの構成情報に基づき（ステップ314）、システム構成の変更を、ディスクサブシステムX及びYにまたがり一括で行う（ステップ315）。

#### 【0034】

ここで本例における構成の変更とは、“特定ホスト10から所定の論理ディスクへのアクセス権限”を、システム管理者Sがその特定ホスト10に対して与えることにある。具体的には、ポート配下のホスト10がアクセス可能な論理ディスクと、当該ポートに接続したホストが装備するホストバスアダプタに付与された、ネットワークで一意的なアドレス、例えば、WWN (World Wide Name) やMACアドレスを、割り当てることを意味する。ここでポートとは、ディスクサブシステム20がホスト10とデータ送受を行う際に用いられる入出力機能をいう。

#### 【0035】

管理サーバ30は、このようなシステム構成の変更を、自己の構成情報データベースへの登録も含め完了し（ステップ316）、占有していた複数のディスクサブシステムX及びYを開放する（ステップ317）。

#### 【0036】

本例のステップ312からステップ317まで、ディスクサブシステムX及びYを排他制御している期間は、他のシステム管理者Tが設定要求を管理サーバ30に対し発行しても、管理サーバ30はシステム管理者Tに対し、システム管理者Sが設定中である旨を通知する。

#### 【0037】

管理サーバ30では、更に、以下の手順で、ディスクサブシステムX及びYの構成情報と、ホスト論理構成情報の対応付けを行う。ここで、ホスト論理構成情報とは、オペレーティングシステム（以下、単にOSと記す）上のファイルと、ファイルを格納した論理ディスクの位置、ファイル容量、データベースと各OSから見た論理ディスクへのアクセスパス情報をいう。

#### 【0038】

尚、各OSから見た論理ディスクへのアクセスパスは、例えば、UNIX系のOSであれば、ホストアダプタカードのIDと、コントローラのIDと論理ディスク番号の3つで特定できる。

#### 【0039】

このような対応付けを行なうのは、システム20の内部の物理的な領域を示すIDと、システム管理者がシステム20を組み込むときに使っているデバイスパスの情報とをリンクさせることで、ホスト10からアクセスするファイルと、このファイルが格納されるディスクサブシステム20の内部の論理ディスクとを対応させ、これらを纏めて管理するためである。

#### 【0040】

##### <ホストエージェントの起動>

複数のホスト10（図1）には、ホストエージェント11がインストールされており、次のイベントに同期してホストエージェント11が起動する。

1) システム管理者の要求によって、管理サーバ30がディスクサブシステム2

0の構成変更が行われ、各ディスクサブシステム20へのシステム構成情報の取得の問合せが為されたとき。

2) ディスクサブシステム20の構成が、障害や保守その他のイベントによって変更され、管理サーバ30がディスクサブシステム20の状態変更を認識し、各ディスクサブシステム20へのシステム構成情報の取得の問合せが為されたとき。

【0041】

ホストエージェント11は、自己を支配するホスト10のOS上の“ホスト論理構成情報”を取得するため、自己のホスト10がアクセス可能であるディスクサブシステム20の論理ディスクに対し、自己のホスト10からその論理ディスクへのアクセスパスを知るためのコマンドを発行する。

【0042】

ホストエージェント11は、OSやデータベース、上位ミドルウェアのアプリケーション・インタフェースを利用して、論理ディスクの内部に格納されているファイルの名称、容量、ファイルが属するファイルシステム上の位置を取得する。

【0043】

管理サーバ30は、各ホストエージェント11が取得した“ホスト論理構成情報”を収集し、ディスクサブシステム20の構成情報に含まれる内部アクセスパスと対応付けて、構成情報データベースに格納する。システム管理者は、管理サーバに問い合わせることで、ファイルが格納されている論理ディスク位置を確認することができる。

【0044】

管理サーバ30は、各ディスクサブシステム20への定期的な問い合わせに同期して、ホストエージェント11が取得する、ホスト10のアプリケーションが使用するファイルの容量のデータを収集し、系時的に管理サーバ30の構成情報データベースに蓄積する。

【0045】

管理サーバ30は、システム管理者がディスクサブシステム20の構成を変更

したときや、障害や保守のイベントで構成が変更されたときも、同様に、システム構成の変更前後の内容を系時的に管理サーバ30の構成情報データベースに蓄積しておく。

#### 【0046】

この結果、システム管理者は、管理サーバ30の構成情報データベースに蓄積された系時的データに対して、時刻、ホスト論理構成情報、システム構成の変更内容をキーとして検索できる。このため、ディスクサブシステムの性能、ファイル容量、その他の諸パラメータについて、ディスクサブシステム20の構成を時間と共に変化させた場合の相関関係を見出し、分析することができる。

#### 【0047】

##### <分析の具体例その1>

以下に具体例を説明する。

第1の具体例は、次に示す事象が発生した場合に、本発明を適用した管理サーバを用いて問題解析を行うケースである。

#### 【0048】

ホスト10において、あるユーザAがデータベースその他のアプリケーションを使用しており、そのファイルの容量を拡張するため、システム管理者が、そのアプリケーションが使用しているディスクサブシステム20において、論理ディスクを追加した場合を考える。この場合、論理ディスクの追加に物理ディスクの追加が伴っても良い。ところが、そのディスクサブシステム20は、別のホストでユーザBが使用する別のアプリケーションB（アプリケーションAよりも、高性能を要求されるものとする）も使用している場合がある。

#### 【0049】

このようなとき、追加された論理ディスクが、高性能を要求されるアプリケーションBが使用するために割り当てられた論理ディスクと、物理資源（物理ディスク）を共有することとなっていた場合には、この論理ディスクの追加が影響し、性能重視のアプリケーションBの実行について、性能低下を引き起こすことになる。従来は、システム管理者が、アプリケーションの性能を監視する場合は、管理サーバから監視が行なえないため、外部の性能監視ツールを使用していた。

この種のツールは、アプリケーションのプロセス稼動状況や、使用するファイルへの読み出しと書き込み性能を監視するものであり、上記のような論理ディスクの追加が、アプリケーションの性能低下を引き起こした原因であることまでは、突き止めることができなかった。

## 【0050】

図4に、アプリケーションの性能がある時点を境に低下した場合に、システム管理者Sが、本発明を適用した管理サーバ30を用いて（構成情報データベースの履歴データを用いて）、執ることができる措置を示す。

## 【0051】

管理サーバ30に、アプリケーションが使用するファイルを特定する情報、そのアプリケーションの性能が低下した時刻を入力する（ステップ411）。入力を受け付けた管理サーバ30は、構成情報データベース内の構成情報テーブルのデータに基づき、ファイルに対応する論理ディスクが、ディスクサブシステム20で、どの物理的領域に位置するかを示す物理ディスク格納位置を特定する。そして、その物理ディスクを共有する他の論理ディスクを検索する（ステップ412）。

## 【0052】

この検索結果を利用して、設定変更の履歴を蓄積しているデータ、例えば、設定変更履歴テーブルの中から、アプリケーションの性能が低下した時刻より前で、物理ディスク格納位置に関連した設定変更の内容を検索する（ステップ413）。

## 【0053】

その設定変更が、アプリケーションの性能低下に関係しているか否かを、管理サーバ30は、論理ディスクの性能履歴を示す、例えば、性能履歴テーブルを参照し、アプリケーションの性能低下時刻の後に、その論理ボリュームの性能が低下しているかを照合し（ステップ414）、低下していれば、該当する設定変更が原因と推定される旨を、システム管理者に通知する（推定原因又は追加時刻）。

## 【0054】

## &lt;分析の具体例その2&gt;

第2の具体例は、次に示す事象が発生した場合に本発明を適用した管理サーバを用いて問題解析を行うケースである。

システム管理者は、アプリケーションが使用するファイル容量の増加傾向を、例えば、四半期毎に定期的に調査し、この増加傾向に対し、ディスクサブシステム20が保有する論理ディスクの容量追加の計画を立てる。この際に、管理サーバが次の手順で計画立案を支援する。

## 【0055】

管理サーバ30は、ホストエージェント11に対して、定期的にファイル容量を問い合わせ、ファイル容量をデータベースに系時的に蓄積する。次に管理サーバ30は、ファイルを格納した論理ディスクとの対応付けを、構成情報データベースの内容から検索し、図5に示すように、その論理ディスクの容量(c4)と、系時的に蓄積したファイル容量と時刻の関係、例えば、開始容量(c1、t1)と、最新容量(c2、t2)のデータをもとに、ファイル容量が論理ディスク容量限度と等しくなる時刻(t4)を予測し、これを論理ディスク追加が必要な時刻、例えば、追加時刻として、システム管理者に通知する。

## 【0056】

システム管理者が、予めファイル容量のしきい値、例えば、ユーザしきい値(c3)を設定しておき、ファイル容量がユーザしきい値を超えたときに(t3時点)、論理ディスクの追加が近い将来に必要となる旨を、システム管理者に警告する。

## 【0057】

更に、管理サーバ30の上位アプリケーションとして、SAN40を形成する構成要素であるホスト10、スイッチ(図示せず)、ディスクサブシステム20の接続形態や、各構成要素の情報を一元管理し、障害監視や性能表示を行う機能を有し、集中管理するアプリケーションがある。この上位アプリケーションは、いちいち、各構成要素に問合せることなく、管理サーバ30に問合せることで、各ディスクサブシステム20の構成や性能の系時的な情報や履歴を、一括して取得できる。



【0058】

【発明の効果】

複数のシステム管理者（又はユーザ）から利用され、かつ、複数のディスクサブシステムの管理を、排他制御コマンドと構成情報データベースの集中管理により行うことができる。

複数のホストから共用される複数のディスクサブシステムの構成で、複数のシステム管理者が、複数のディスクサブシステムにまたがったシステム構成を、一括して定義できる効果がある。

システム全体の構成情報を系時的に蓄積した構成情報データベースを用いて、システム構成の履歴管理を行なうことができる。

システム構成の変更が、アプリケーションに及ぼす影響を、正確に見積もることができる効果がある。アプリケーションの性能低下の原因究明が可能である。

システム構成の変更の時期や、論理ディスク容量の追加の時期を、計画し、予測して通知できる効果がある。

【図面の簡単な説明】

【図1】

複数のホストコンピュータが、ネットワークを介して、データを送受し共用する複数のディスクサブシステムと構成するシステムの概略を示す図である。

【図2】

本発明を適用する管理サーバの機能ブロックを示す図である。

【図3】

複数のディスクサブシステムにまたがった所望の設定を一括して行なう手順の流れを例示するための図である。

【図4】

本発明を用いた分析の一例であって、アプリケーションの性能低下を、管理サーバが所有する構成情報データベースを用いて、構成情報の履歴を追跡し、原因を究明する流れを示した図である。

【図5】

本発明を用いた分析の一例であって、論理ディスクの増設計画を管理サーバ所

有のファイル容量の履歴を利用して分析する流れを示した図である。

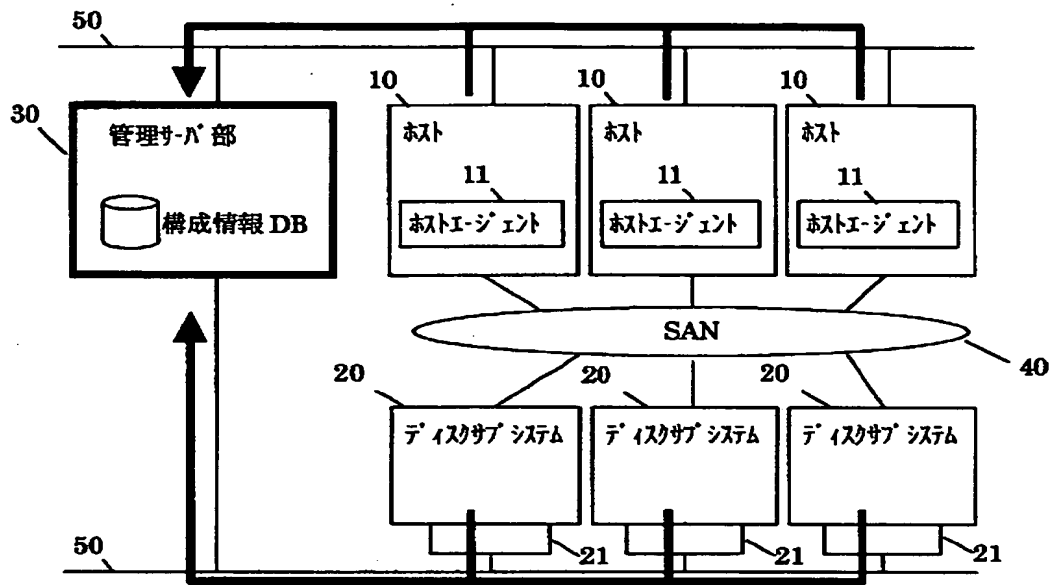
【符号の説明】

10・・・ホスト	11・・・ホストエージェント
20・・・ディスクサブシステム	21・・・外部接続用インタフェース
30・・・管理サーバ	31・・・ユーザ管理レイヤ
32・・・オブジェクト管理レイヤ	33・・・エージェント管理レイヤ
34・・・インタフェースレイヤ	
341・・・サブシステムインタフェース	
342・・・ホストインタフェース	
40・・・ストレージエリアネットワーク (SAN)	
50・・・ローカルエリアネットワーク (LAN)	

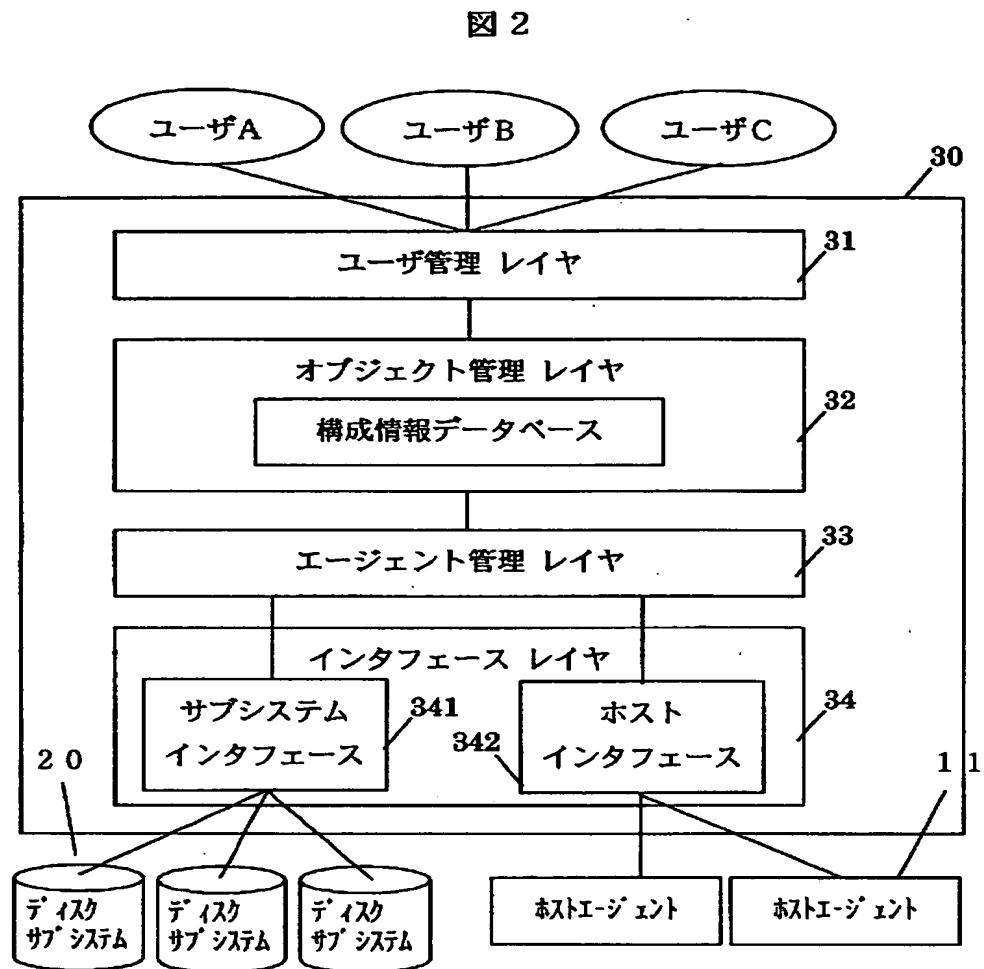
【書類名】 図面

【図1】

図 1

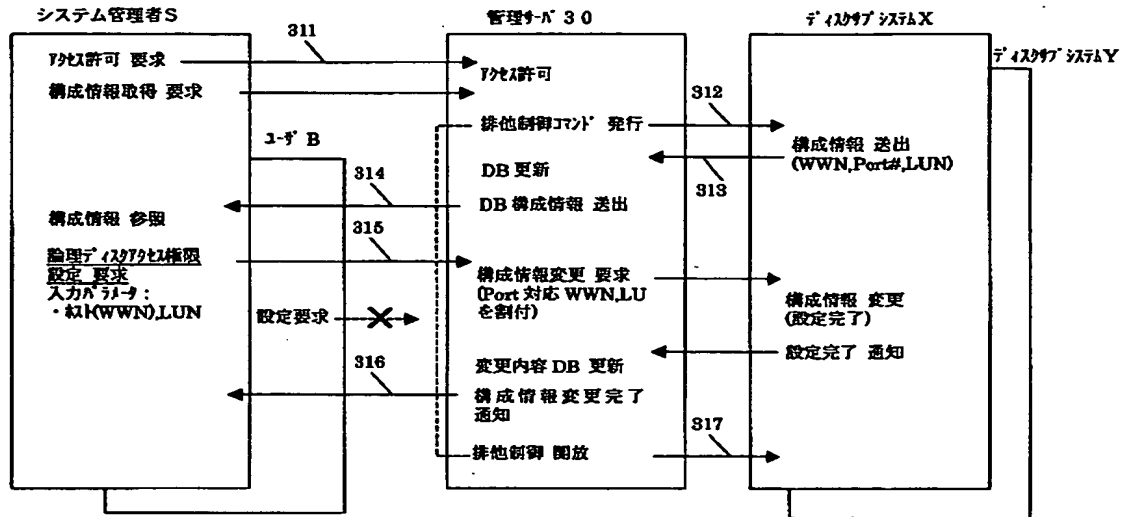


【図2】



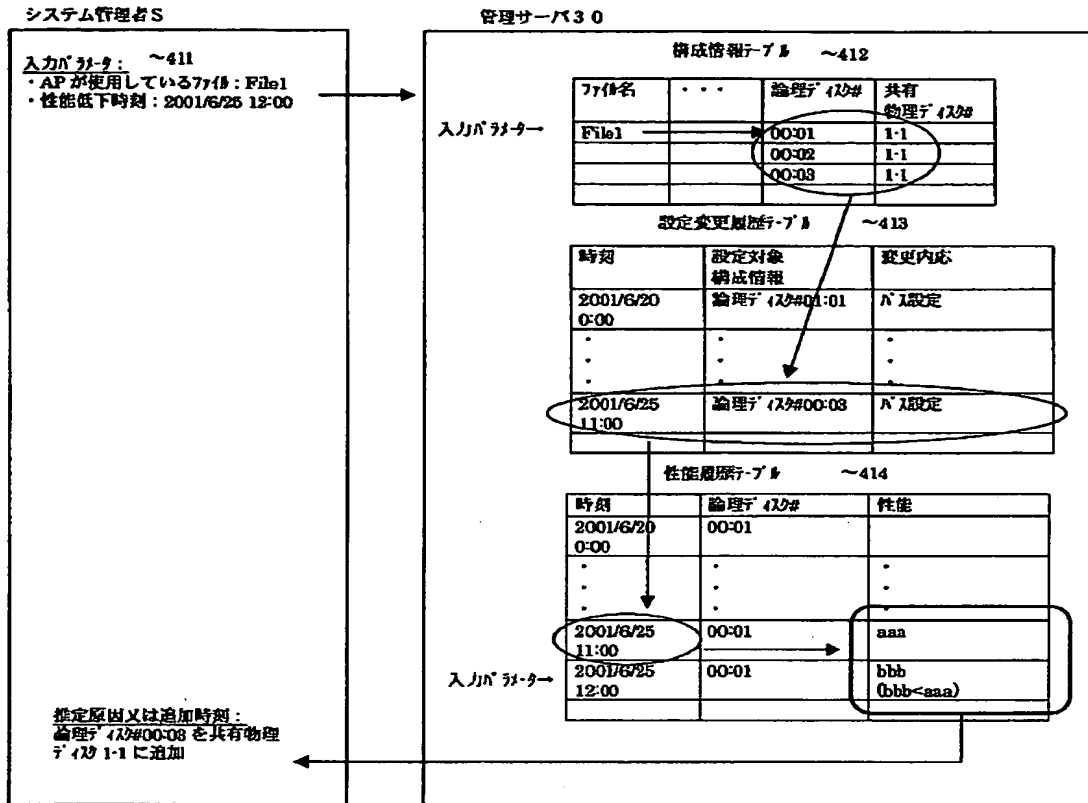
【図 3】

図 3



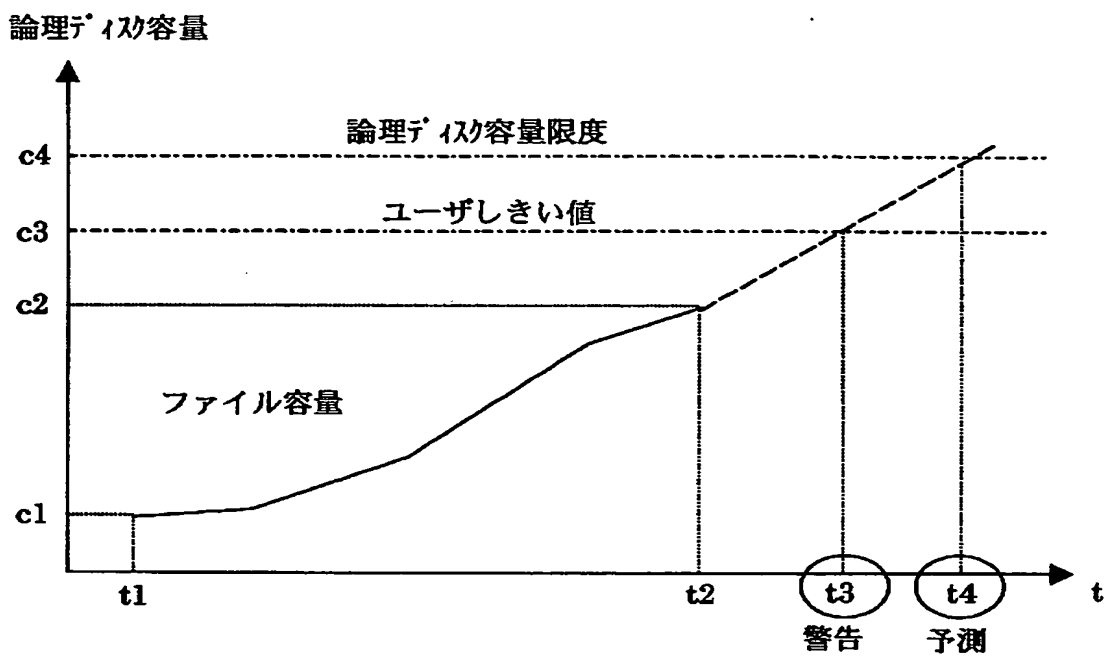
【図 4】

図 4



【図5】

図 5



【書類名】 要約書

【要約】

【課題】

複数（N台）のホストから共用される、複数（M台）のディスクサブシステムの構成において、複数のシステム管理者（L人）が、M台のディスクサブシステムを横断的に一括して管理し、所定の設定を迅速かつ簡易に実現する。システムの構成の変更が、ホストで実行されるアプリケーションの性能に与える影響を把握可能とする。システムにおける計画的な論理ディスク容量の追加時期の決定を可能とする。

【解決手段】

1) 一時的に全ての複数のディスクサブシステムへのアクセスを制限する排他制御コマンドを設ける。2) この排他制御コマンドにより、複数のディスクサブシステム全体の構成情報、例えば、性能、設定の変更を、系時的に取得し、管理サーバのデータベース（“構成情報データベース”）に格納して一元管理する。3) アプリケーションが使用しているファイルの、論理ディスク上の位置を検出する機能を用いて、“構成情報データベース”に対応付けする機能を設ける。4) “構成情報データベース”に、システム構成の変更内容や時刻をキーに検索できる手段を設ける。

ストレージシステム全体の構成情報を系時的に一元管理できるので、システム構成の変更に起因するアプリケーション性能の低下や、ストレージシステムへの計画的な論理ディスク容量の追加、その他イベントの分析・予測が容易に行える効果がある。

【選択図】 図2



認定・付加情報

特許出願の番号	特願2001-295397
受付番号	50101421715
書類名	特許願
担当官	第七担当上席 0096
作成日	平成13年 9月28日

<認定情報・付加情報>

【提出日】	平成13年 9月27日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所